

SEQUENCE LISTING

<110> Burgin, Alex B.  
Stewart, Lance J.

<120> Use of Phosphorothiolate Polynucleotides  
In Ligating Nucleic Acids

<130> P-RD 4806

<150> US 60/290,313

<151> 2001-05-10

<160> 19

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3739

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (247)...(2541)

<400> 1

```

caaatgCGAA cttaggtgt tacacaactg ctgggggtctg ttctcgccgc cgcgccgga 60
gtcaggcagc gtcgccgcg tggtagcagc ctacagcgtt tctggagtct cgggccaca 120
gtcaccgcgc cttacctgcg cctcctcgag cctccggagt ccccgccgc cgcacaggg 180
cggttcgccg ttctcgcttc ccccacgcgc cctcgccgc cgcgcgcgc gtccctccg 240
gccgac atg agt ggg gac cac ctc cac aac gat tcc cag atc gaa gcg 288
Met Ser Gly Asp His Leu His Asn Asp Ser Gln Ile Glu Ala
1 5 10

```

```

gat ttc cga ttg aat gat tct cat aaa cac aaa gat aaa cac aaa gat 336
Asp Phe Arg Leu Asn Asp Ser His Lys His Lys Asp Lys His Lys Asp
15 20 25 30

```

```

cga gaa cac cgg cac aaa gaa cac aag aag gag aag gac cgg gaa aag 384
Arg Glu His Arg His Lys Glu His Lys Lys Glu Lys Asp Arg Glu Lys
35 40 45

```

```

tcc aag cat agc aac agt gaa cat aaa gat tct gaa aag aaa cac aaa 432
Ser Lys His Ser Asn Ser Glu His Lys Asp Ser Glu Lys Lys His Lys
50 55 60

```

```

gag aag gag aag acc aaa cac aaa gat gga agc tca gaa aag cat aaa 480
Glu Lys Glu Lys Thr Lys His Lys Asp Gly Ser Ser Glu Lys His Lys
65 70 75

```

gac aaa cat aaa gac aga gac aag gaa aaa cga aaa gag gaa aag gtt Asp Lys His Lys Asp Arg Asp Lys Glu Lys Arg Lys Glu Glu Lys Val 80 85 90	528
cga gcc tct ggg gat gca aaa ata aag aag gag aag gaa aat ggc ttc Arg Ala Ser Gly Asp Ala Lys Ile Lys Lys Glu Lys Glu Asn Gly Phe 95 100 105 110	576
tct agt cca cca caa att aaa gat gaa cct gaa gat gat ggc tat ttt Ser Ser Pro Pro Gln Ile Lys Asp Glu Pro Glu Asp Asp Gly Tyr Phe 115 120 125	624
gtt cct cct aaa gag gat ata aag cca tta aag aga cct cga gat gag Val Pro Pro Lys Glu Asp Ile Lys Pro Leu Lys Arg Pro Arg Asp Glu 130 135 140	672
gat gat gct gat tat aaa cct aag aaa att aaa aca gaa gat acc aag Asp Asp Ala Asp Tyr Lys Pro Lys Lys Ile Lys Thr Glu Asp Thr Lys 145 150 155	720
aag gag aag aaa aga aaa cta gaa gaa gaa gag gat ggt aaa ttg aaa Lys Glu Lys Lys Arg Lys Leu Glu Glu Glu Asp Gly Lys Leu Lys 160 165 170	768
aaa ccc aag aat aaa gat aaa gat aaa aaa gtt cct gag cca gat aac Lys Pro Lys Asn Lys Asp Lys Asp Lys Lys Val Pro Glu Pro Asp Asn 175 180 185 190	816
aag aaa aag aag ccg aag aaa gaa gag gaa cag aag tgg aaa tgg tgg Lys Lys Lys Lys Pro Lys Lys Glu Glu Glu Gln Lys Trp Lys Trp Trp 195 200 205	864
gaa gaa gag cgc tat cct gaa ggc atc aag tgg aaa ttc cta gaa cat Glu Glu Glu Arg Tyr Pro Glu Gly Ile Lys Trp Lys Phe Leu Glu His 210 215 220	912
aaa ggt cca gta ttt gcc cca cca tat gag cct ctt cca gag aat gtc Lys Gly Pro Val Phe Ala Pro Pro Tyr Glu Pro Leu Pro Glu Asn Val 225 230 235	960
aag ttt tat tat gat ggt aaa gtc atg aag ctg agc ccc aaa gca gag Lys Phe Tyr Tyr Asp Gly Lys Val Met Lys Leu Ser Pro Lys Ala Glu 240 245 250	1008
gaa gta gct acg ttc ttt gca aaa atg ctc gac cat gaa tat act acc Glu Val Ala Thr Phe Phe Ala Lys Met Leu Asp His Glu Tyr Thr Thr 255 260 265 270	1056
aag gaa ata ttt agg aaa aat ttc ttt aaa gac tgg aga aag gaa atg Lys Glu Ile Phe Arg Lys Asn Phe Phe Lys Asp Trp Arg Lys Glu Met 275 280 285	1104
act aat gaa gag aag aat att atc acc aac cta agc aaa tgt gat ttt	1152

Thr Asn Glu Glu Lys Asn Ile Ile Thr Asn Leu Ser Lys Cys Asp Phe	1200
290 295 300	
acc cag atg agc cag tat ttc aaa gcc cag acg gaa gct cgg aaa cag	1200
Thr Gln Met Ser Gln Tyr Phe Lys Ala Gln Thr Glu Ala Arg Lys Gln	
305 310 315	
atg agc aag gaa gag aaa ctg aaa atc aaa gag gag aat gaa aaa tta	1248
Met Ser Lys Glu Glu Lys Leu Lys Ile Lys Glu Glu Asn Glu Lys Leu	
320 325 330	
ctg aaa gaa tat gga ttc tgt att atg gat aac cac aaa gag agg att	1296
Leu Lys Glu Tyr Gly Phe Cys Ile Met Asp Asn His Lys Glu Arg Ile	
335 340 345 350	
gct aac ttc aag ata gag cct cct gga ctt ttc cgt gcc cgc gcc aac	1344
Ala Asn Phe Lys Ile Glu Pro Pro Gly Leu Phe Arg Gly Arg Gly Asn	
355 360 365	
cac ccc aag atg gcc atg ctg aag aga cga atc atg ccc gag gat ata	1392
His Pro Lys Met Gly Met Leu Lys Arg Arg Ile Met Pro Glu Asp Ile	
370 375 380	
atc atc aac tgt agc aaa gat gcc aag gtt cct tct cct cct cca gga	1440
Ile Ile Asn Cys Ser Lys Asp Ala Lys Val Pro Ser Pro Pro Gly	
385 390 395	
cat aag tgg aaa gaa gtc cgg cat gat aac aag gtt act tgg ctg gtt	1488
His Lys Trp Lys Glu Val Arg His Asp Asn Lys Val Thr Trp Leu Val	
400 405 410	
tcc tgg aca gag aac atc caa ggt tcc att aaa tac atc atg ctt aac	1536
Ser Trp Thr Glu Asn Ile Gln Gly Ser Ile Lys Tyr Ile Met Leu Asn	
415 420 425 430	
cct agt tca cga atc aag ggt gag aag gac tgg cag aaa tac gag act	1584
Pro Ser Ser Arg Ile Lys Gly Glu Lys Asp Trp Gln Lys Tyr Glu Thr	
435 440 445	
gct cgg cgg ctg aaa aaa tgt gtg gac aag atc cgg aac cag tat cga	1632
Ala Arg Arg Leu Lys Lys Cys Val Asp Lys Ile Arg Asn Gln Tyr Arg	
450 455 460	
gaa gac tgg aag tcc aaa gag atg aaa gtc cgg cag aga gct gta gcc	1680
Glu Asp Trp Lys Ser Lys Glu Met Lys Val Arg Gln Arg Ala Val Ala	
465 470 475	
ctg tac ttc atc gac aag ctt gct ctg aga gca gcc aat gaa aag gag	1728
Leu Tyr Phe Ile Asp Lys Leu Ala Leu Arg Ala Gly Asn Glu Lys Glu	
480 485 490	
gaa gga gaa aca cgg gac act gtg gcc tgc tgc tca ctt cgt gtg gag	1776
Glu Gly Glu Thr Ala Asp Thr Val Gly Cys Cys Ser Leu Arg Val Glu	

495	500	505	510	
cac atc aat cta cac cca gag ttg gat ggt cag gaa tat gtg gta gag	His Ile Asn Leu	His Pro Glu Leu Asp Gly Gln Glu Tyr Val Val Glu	1824	
	515	520	525	
ttt gac ttc ctc ggg aag gac tcc atc aga tac tat aac aag gtc cct	Phe Asp Phe Leu Gly Lys Asp Ser	Ile Arg Tyr Tyr Asn Lys Val Pro	1872	
	530	535	540	
gtt gag aaa cga gtt ttt aag aac cta caa cta ttt atg gag aac aag	Val Glu Lys Arg Val Phe Lys Asn Leu Gln Leu Phe Met Glu Asn Lys		1920	
	545	550	555	
cag ccc gag gat gat ctt ttt gat aga ctc aat act ggt att ctg aat	Gln Pro Glu Asp Asp Leu Phe Asp Arg Leu Asn Thr Gly Ile Leu Asn		1968	
	560	565	570	
aag cat ctt cag gat ctc atg gag ggc ttg aca gcc aag gta ttc cgt	Lys His Leu Gln Asp Leu Met Glu Gly Leu Thr Ala Lys Val Phe Arg		2016	
	575	580	585	590
aca tac aat gcc tcc atc acg cta cag cag cag cta aaa gaa ctg aca	Thr Tyr Asn Ala Ser Ile Thr Leu Gln Gln Leu Lys Glu Leu Thr		2064	
	595	600	605	
gcc ccg gat gag aac atc cca gcg aag atc ctt tct tat aac cgt gcc	Ala Pro Asp Glu Asn Ile Pro Ala Lys Ile Leu Ser Tyr Asn Arg Ala		2112	
	610	615	620	
aat cga gct gtt gca att ctt tgt aac cat cag agg gca cca cca aaa	Asn Arg Ala Val Ala Ile Leu Cys Asn His Gln Arg Ala Pro Pro Lys		2160	
	625	630	635	
act ttt gag aag tct atg atg aac ttg caa act aag att gat gcc aag	Thr Phe Glu Lys Ser Met Met Asn Leu Gln Thr Lys Ile Asp Ala Lys		2208	
	640	645	650	
aag gaa cag cta gca gat gcc cgg aga gac ctg aaa agt gct aag gct	Lys Glu Gln Leu Ala Asp Ala Arg Arg Asp Leu Lys Ser Ala Lys Ala		2256	
	655	660	665	670
gat gcc aag gtc atg aag gat gca aag acg aag aag gta gta gag tca	Asp Ala Lys Val Met Lys Asp Ala Lys Thr Lys Lys Val Val Glu Ser		2304	
	675	680	685	
aag aag aag gct gtt cag aga ctg gag gaa cag ttg atg aag ctg gaa	Lys Lys Lys Ala Val Gln Arg Leu Glu Glu Gln Leu Met Lys Leu Glu		2352	
	690	695	700	
gtt caa gcc aca gac cga gag gaa aat aaa cag att gcc ctg gga acc	Val Gln Ala Thr Asp Arg Glu Glu Asn Lys Gln Ile Ala Leu Gly Thr		2400	
	705	710	715	

tcc aaa ctc aat tat ctg gac cct agg atc aca gtg gct tgg tgc aag 2448  
Ser Lys Leu Asn Tyr Leu Asp Pro Arg Ile Thr Val Ala Trp Cys Lys  
720 725 730

aag tgg ggt gtc cca att gag aag att tac aac aaa acc cag cgg gag 2496  
Lys Trp Gly Val Pro Ile Glu Lys Ile Tyr Asn Lys Thr Gln Arg Glu  
735 740 745 750

aag ttt gcc tgg gcc att gac atg gct gat gaa gac tat gag ttt 2541  
Lys Phe Ala Trp Ala Ile Asp Met Ala Asp Glu Asp Tyr Glu Phe  
755 760 765

tagccagtct caagaggcag agttctgtga agaggaaacag tgtggtttgg gaaagatgga 2601  
taaaactgagc ctcaactggc ctctgtgcctg ggggagagag gcagcaagtc ttaacaaaacc 2661  
aacatcttttg cgaaaagata aacctggaga tattataagg gagagctgag ccagttgtcc 2721  
tatggacaac ttattttaaaa atatttcaga tatcaaaatt ctagctgtat gattttgttt 2781  
gaattttgtt ttatttttca agagggcgaag tggatgggaa ttgtcagcg ttctaccagg 2841  
caaattcact gtttccactga aatgttttga ttctcttagc tactgtatgc aaagtccgat 2901  
tatattgggt cgtttttaca gttagggttt tgcaataact tctatatttt aatagaataa 2961  
aattcctaata ctcccttccc tctctcccat ttccaggaatt taaaattaag tagaacaata 3021  
aaccagcgc acctgttaga gtctgcactc tctattgtca tggggatcaa ttctattaa 3081  
acctgaagca gtcgtggctt tggcagtggt ttggttcaga cacctgttca cagaaaaagc 3141  
atgatgggaa aatatttctc gactttgagtg ttctttttaa aatgtgaatt ttattttct 3201  
tttaattatt ttaaaatatt taaacctttt tctgtatctt aaagatcggt tagattgggg 3261  
tgggggaggg atgaagggcg agtgaatcta aggataatga aataatcagt gactgaaacc 3321  
attttcccat catctttgt tctgagcatt cgctgtacc tttaagatat ccatcttttt 3381  
ctttttaacc taatctttc acttgaaga ttttattgta taaaaagttt cacaggtcaa 3441  
taaaactaga ggaataatgag tattttggtc aaaaaaagga aaaataatca agattttagg 3501  
gctttttatt ttctttttgt aattgtgtaa aaaatggaaa aaacacataaa aagcagaatt 3561  
ttaattgtga gacatttttt gctataatca ttagttttag aggcattgtt agtttagtgt 3621  
gtgtgcagag tccattttcc acatctttcc tcaagtatct tctattttta tcatgaaacc 3681  
ccttttaac aactgtaggt tatttaaaat aaattcctac aacttaatgg aaacttaa 3739

<210> 2  
<211> 765  
<212> PRT  
<213> Homo sapiens

<400> 2  
Met Ser Gly Asp His Leu His Asn Asp Ser Gln Ile Glu Ala Asp Phe  
1 5 10 15  
Arg Leu Asn Asp Ser His Lys His Lys Asp Lys His Lys Asp Arg Glu  
20 25 30  
His Arg His Lys Glu His Lys Lys Glu Lys Asp Arg Glu Lys Ser Lys  
35 40 45  
His Ser Asn Ser Glu His Lys Asp Ser Glu Lys Lys His Lys Glu Lys  
50 55 60  
Glu Lys Thr Lys His Lys Asp Gly Ser Ser Glu Lys His Lys Asp Lys  
65 70 75 80  
His Lys Asp Arg Asp Lys Glu Lys Arg Lys Glu Lys Val Arg Ala  
85 90 95  
Ser Gly Asp Ala Lys Ile Lys Lys Glu Lys Glu Asn Gly Phe Ser Ser

[illegible]

Phe Leu Gly Lys Asp Ser Ile Arg Tyr Tyr Asn Lys Val Pro Val Glu  
530 535 540  
Lys Arg Val Phe Lys Asn Leu Gln Leu Phe Met Glu Asn Lys Gln Pro  
545 550 555 560  
Glu Asp Asp Leu Phe Asp Arg Leu Asn Thr Gly Ile Leu Asn Lys His  
565 570 575  
Leu Gln Asp Leu Met Glu Gly Leu Thr Ala Lys Val Phe Arg Thr Tyr  
580 585 590  
Asn Ala Ser Ile Thr Leu Gln Gln Gln Leu Lys Glu Leu Thr Ala Pro  
595 600 605  
Asp Glu Asn Ile Pro Ala Lys Ile Leu Ser Tyr Asn Arg Ala Asn Arg  
610 615 620  
Ala Val Ala Ile Leu Cys Asn His Gln Arg Ala Pro Pro Lys Thr Phe  
625 630 635 640  
Glu Lys Ser Met Met Asn Leu Gln Thr Lys Ile Asp Ala Lys Lys Glu  
645 650 655  
Gln Leu Ala Asp Ala Arg Arg Asp Leu Lys Ser Ala Lys Ala Asp Ala  
660 665 670  
Lys Val Met Lys Asp Ala Lys Thr Lys Lys Val Val Glu Ser Lys Lys  
675 680 685  
Lys Ala Val Gln Arg Leu Glu Glu Gln Leu Met Lys Leu Glu Val Gln  
690 695 700  
Ala Thr Asp Arg Glu Glu Asn Lys Gln Ile Ala Leu Gly Thr Ser Lys  
705 710 715 720  
Leu Asn Tyr Leu Asp Pro Arg Ile Thr Val Ala Trp Cys Lys Lys Trp  
725 730 735  
Gly Val Pro Ile Glu Lys Ile Tyr Asn Lys Thr Gln Arg Glu Lys Phe  
740 745 750  
Ala Trp Ala Ile Asp Met Ala Asp Glu Asp Tyr Glu Phe  
755 760 765

<210> 3  
<211> 1719  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> human fusion

<221> CDS  
<222> (1)... (1716)

<400> 3  
atg aac aag aaa aag aag ccg aag aaa gaa gag cat cat cat cat cat 48  
Met Asn Lys Lys Lys Lys Pro Lys Lys Glu His His His His His  
1 5 10 15

cac atc gaa ggt cgt atc aag tgg aaa ttc cta gaa cat aaa ggt cca 96  
His Ile Glu Gly Arg Ile Lys Trp Lys Phe Leu Glu His Lys Gly Pro  
20 25 30

gta ttt gcc cca cca tat gag cct ctt cca gag aat gtc aag ttt tat 144

Val Phe Ala Pro Pro Tyr Glu Pro Leu Pro Glu Asn Val Lys Phe Tyr	
35 40 45	
tat gat ggt aaa gtc atg aag ctg agc ccc aaa gca gag gaa gta gct	192
Tyr Asp Gly Lys Val Met Lys Leu Ser Pro Lys Ala Glu Glu Val Ala	
50 55 60	
acg ttc ttt gca aaa atg ctc gac cat gaa tat act acc aag gaa ata	240
Thr Phe Phe Ala Lys Met Leu Asp His Glu Tyr Thr Thr Lys Glu Ile	
65 70 75 80	
ttt agg aaa aat ttc ttt aaa gac tgg aga aag gaa atg act aat gaa	288
Phe Arg Lys Asn Phe Phe Lys Asp Trp Arg Lys Glu Met Thr Asn Glu	
85 90 95	
gag aag aat att atc acc aac cta agc aaa tgt gat ttt acc cag atg	336
Glu Lys Asn Ile Ile Thr Asn Leu Ser Lys Cys Asp Phe Thr Gln Met	
100 105 110	
agc cag tat ttc aaa gcc cag acg gaa gct cgg aaa cag atg agc aag	384
Ser Gln Tyr Phe Lys Ala Gln Thr Glu Ala Arg Lys Gln Met Ser Lys	
115 120 125	
gaa gag aaa ctg aaa atc aaa gag gag aat gaa aaa tta ctg aaa gaa	432
Glu Glu Lys Leu Lys Ile Lys Glu Glu Asn Glu Lys Leu Leu Lys Glu	
130 135 140	
tat gga ttc tgt att atg gat aac cac aaa gag agg att get aac ttc	480
Tyr Gly Phe Cys Ile Met Asp Asn His Lys Glu Arg Ile Ala Asn Phe	
145 150 155 160	
aag ata gag cct cct gga ctt ttc cgt gcc cgc gcc aac cac ccc aag	528
Lys Ile Glu Pro Pro Gly Leu Phe Arg Gly Arg Gly Asn His Pro Lys	
165 170 175	
atg gcc atg ctg aag aga cga atc atg ccc gag gat ata atc atc aac	576
Met Gly Met Leu Lys Arg Arg Ile Met Pro Glu Asp Ile Ile Ile Asn	
180 185 190	
tgt agc aaa gat gcc aag gtt cct tct cct cct cca gga cat aag tgg	624
Cys Ser Lys Asp Ala Lys Val Pro Ser Pro Pro Pro Gly His Lys Trp	
195 200 205	
aaa gaa gtc cgg cat gat aac aag gtt act tgg ctg gtt tcc tgg aca	672
Lys Glu Val Arg His Asp Asn Lys Val Thr Trp Leu Val Ser Trp Thr	
210 215 220	
gag aac atc caa ggt tcc att aaa tac atc atg ctt aac cct agt tca	720
Glu Asn Ile Gln Gly Ser Ile Lys Tyr Ile Met Leu Asn Pro Ser Ser	
225 230 235 240	
cga atc aag ggt gag aag gac tgg cag aaa tac gag act gct cgg cgg	768
Arg Ile Lys Gly Glu Lys Asp Trp Gln Lys Tyr Glu Thr Ala Arg Arg	



245	250	255	
ctg aaa aaa tgt gtg gac aag atc Leu Lys Lys Cys Val Asp Lys Ile	cgg aac cag tat cga gaa gac tgg Arg Asn Gln Tyr Arg Glu Asp Trp		816
260	265	270	
aag tcc aaa gag atg aaa gtc cgg cag aga gct gta gcc ctg tac ttc Lys Ser Lys Glu Met Lys Val Arg Gln Arg Ala Val Ala Leu Tyr Phe			864
275	280	285	
atc gac aag ctt gct ctg aga gca ggc aat gaa aag gag gaa gga gaa Ile Asp Lys Leu Ala Leu Arg Ala Gly Asn Glu Lys Glu Glu Gly Glu			912
290	295	300	
aca gcg gac act gtg ggc tgc tgc tca ctt cgt gtg gag cac atc aat Thr Ala Asp Thr Val Gly Cys Cys Ser Leu Arg Val Glu His Ile Asn			960
305	310	315	320
cta cac cca gag ttg gat ggt cag gaa tat gtg gta gag ttt gac ttc Leu His Pro Glu Leu Asp Gly Gln Glu Tyr Val Val Glu Phe Asp Phe			1008
325	330	335	
ctc ggg aag gac tcc atc aga tac tat aac aag gtc cct gtt gag aaa Leu Gly Lys Asp Ser Ile Arg Tyr Tyr Asn Lys Val Pro Val Glu Lys			1056
340	345	350	
cga gtt ttt aag aac cta caa cta ttt atg gag aac aag cag ccc gag Arg Val Phe Lys Asn Leu Gln Leu Phe Met Glu Asn Lys Gln Pro Glu			1104
355	360	365	
gat gat ctt ttt gat aga ctc aat act ggt att ctg aat aag cat ctt Asp Asp Leu Phe Asp Arg Leu Asn Thr Gly Ile Leu Asn Lys His Leu			1152
370	375	380	
cag gat ctc atg gag ggc ttg aca gcc aag gta ttc cgt acg tac aat Gln Asp Leu Met Glu Gly Leu Thr Ala Lys Val Phe Arg Thr Tyr Asn			1200
385	390	395	400
gcc tcc atc acg cta cag cag cag cta aaa gaa ctg aca gcc ccg gat Ala Ser Ile Thr Leu Gln Gln Gln Leu Lys Glu Leu Thr Ala Pro Asp			1248
405	410	415	
gag aac atc cca gcg aag atc ctt tct tat aac cgt gcc aat cga gct Glu Asn Ile Pro Ala Lys Ile Leu Ser Tyr Asn Arg Ala Asn Arg Ala			1296
420	425	430	
gtt gca att ctt tgt aac cat cag agg gca cca cca aaa act ttt gag Val Ala Ile Leu Cys Asn His Gln Arg Ala Pro Pro Lys Thr Phe Glu			1344
435	440	445	
aag tct atg atg aac ttg caa act aag att gat gcc aag aag gaa cag Lys Ser Met Met Asn Leu Gln Thr Lys Ile Asp Ala Lys Lys Glu Gln			1392
450	455	460	

cta gca gat gcc cgg aga gac ctg aaa agt gct aag gct gat gcc aag 1440  
Leu Ala Asp Ala Arg Arg Asp Leu Lys Ser Ala Lys Ala Asp Ala Lys  
465 470 475 480

gtc atg aag gat gca aag acg aag aag gta gta gag tca aag aag aag 1488  
Val Met Lys Asp Ala Lys Thr Lys Lys Val Val Glu Ser Lys Lys Lys  
485 490 495

gct gtt cag aga ctg gag gaa cag ttg atg aag ctg gaa gtt caa gcc 1536  
Ala Val Gln Arg Leu Glu Glu Gln Leu Met Lys Leu Glu Val Gln Ala  
500 505 510

aca gac cga gag gaa aat aaa cag att gcc ctg gga acc tcc aaa ctc 1584  
Thr Asp Arg Glu Glu Asn Lys Gln Ile Ala Leu Gly Thr Ser Lys Leu  
515 520 525

aat tat ctg gac cct agg atc aca gtg gct tgg tgc aag aag tgg ggt 1632  
Asn Tyr Leu Asp Pro Arg Ile Thr Val Ala Trp Cys Lys Lys Trp Gly  
530 535 540

gtc cca att gag aag att tac aac aaa acc cag cgg gag aag ttt gcc 1680  
Val Pro Ile Glu Lys Ile Tyr Asn Lys Thr Gln Arg Glu Lys Phe Ala  
545 550 555 560

tgg gcc att gac atg gct gat gaa gac tat gag ttt tag 1719  
Trp Ala Ile Asp Met Ala Asp Gln Asp Tyr Glu Phe  
565 570

<210> 4  
<211> 572  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> human fusion

<400> 4  
Met Asn Lys Lys Lys Lys Pro Lys Lys Glu Glu His His His His His  
1 5 10 15  
His Ile Glu Gly Arg Ile Lys Trp Lys Phe Leu Glu His Lys Gly Pro  
20 25 30  
Val Phe Ala Pro Pro Tyr Glu Pro Leu Pro Glu Asn Val Lys Phe Tyr  
35 40 45  
Tyr Asp Gly Lys Val Met Lys Leu Ser Pro Lys Ala Glu Glu Val Ala  
50 55 60  
Thr Phe Phe Ala Lys Met Leu Asp His Glu Tyr Thr Thr Lys Glu Ile  
65 70 75 80  
Phe Arg Lys Asn Phe Phe Lys Asp Trp Arg Lys Glu Met Thr Asn Glu  
85 90 95  
Glu Lys Asn Ile Ile Thr Asn Leu Ser Lys Cys Asp Phe Thr Gln Met  
100 105 110

Ser Gln Tyr Phe Lys Ala Gln Thr Glu Ala Arg Lys Gln Met Ser Lys  
115 120 125  
Glu Glu Lys Leu Lys Ile Lys Glu Glu Asn Glu Lys Leu Leu Lys Glu  
130 135 140  
Tyr Gly Phe Cys Ile Met Asp Asn His Lys Glu Arg Ile Ala Asn Phe  
145 150 155 160  
Lys Ile Glu Pro Pro Gly Leu Phe Arg Gly Arg Gly Asn His Pro Lys  
165 170 175  
Met Gly Met Leu Lys Arg Arg Ile Met Pro Glu Asp Ile Ile Ile Asn  
180 185 190  
Cys Ser Lys Asp Ala Lys Val Pro Ser Pro Pro Gly His Lys Trp  
195 200 205  
Lys Glu Val Arg His Asp Asn Lys Val Thr Trp Leu Val Ser Trp Thr  
210 215 220  
Glu Asn Ile Gln Gly Ser Ile Lys Tyr Ile Met Leu Asn Pro Ser Ser  
225 230 235 240  
Arg Ile Lys Gly Glu Lys Asp Trp Gln Lys Tyr Glu Thr Ala Arg Arg  
245 250 255  
Leu Lys Lys Cys Val Asp Lys Ile Arg Asn Gln Tyr Arg Glu Asp Trp  
260 265 270  
Lys Ser Lys Glu Met Lys Val Arg Gln Arg Ala Val Ala Leu Tyr Phe  
275 280 285  
Ile Asp Lys Leu Ala Leu Arg Ala Gly Asn Glu Lys Glu Glu Gly Glu  
290 295 300  
Thr Ala Asp Thr Val Gly Cys Cys Ser Leu Arg Val Glu His Ile Asn  
305 310 315 320  
Leu His Pro Glu Leu Asp Gly Gln Glu Tyr Val Val Glu Phe Asp Phe  
325 330 335  
Leu Gly Lys Asp Ser Ile Arg Tyr Tyr Asn Lys Val Pro Val Glu Lys  
340 345 350  
Arg Val Phe Lys Asn Leu Gln Leu Phe Met Glu Asn Lys Gln Pro Glu  
355 360 365  
Asp Asp Leu Phe Asp Arg Leu Asn Thr Gly Ile Leu Asn Lys His Leu  
370 375 380  
Gln Asp Leu Met Glu Gly Leu Thr Ala Lys Val Phe Arg Thr Tyr Asn  
385 390 395 400  
Ala Ser Ile Thr Leu Gln Gln Gln Leu Lys Glu Leu Thr Ala Pro Asp  
405 410 415  
Glu Asn Ile Pro Ala Lys Ile Leu Ser Tyr Asn Arg Ala Asn Arg Ala  
420 425 430  
Val Ala Ile Leu Cys Asn His Gln Arg Ala Pro Pro Lys Thr Phe Glu  
435 440 445  
Lys Ser Met Met Asn Leu Gln Thr Lys Ile Asp Ala Lys Lys Glu Gln  
450 455 460  
Leu Ala Asp Ala Arg Arg Asp Leu Lys Ser Ala Lys Ala Asp Ala Lys  
465 470 475 480  
Val Met Lys Asp Ala Lys Thr Lys Lys Val Val Glu Ser Lys Lys Lys  
485 490 495  
Ala Val Gln Arg Leu Glu Glu Gln Leu Met Lys Leu Glu Val Gln Ala  
500 505 510  
Thr Asp Arg Glu Glu Asn Lys Gln Ile Ala Leu Gly Thr Ser Lys Leu  
515 520 525  
Asn Tyr Leu Asp Pro Arg Ile Thr Val Ala Trp Cys Lys Lys Trp Gly

530	535	540
Val Pro Ile Glu Lys Ile Tyr Asn Lys Thr Gln Arg Glu Lys Phe Ala		
545	550	555
Trp Ala Ile Asp Met Ala Asp Glu Asp Tyr Glu Phe		560
	565	570

<210> 5  
<211> 6  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic

<400> 5  
gycctt

6

<210> 6  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (18)...(19)  
<223> n=5-bromo-deoxyuridine

<223> synthetic

<400> 6  
aaaaagactt agaaaaannt tt

22

<210> 7  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<221> modified\_base  
<222> (1)...(22)  
<223> n=5-iodo-deoxyuridine

<223> synthetic

<400> 7  
aaaaagacnn tgaaaaannn nt

22

<210> 8  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic

<400> 8  
aatcgcggc cgcaaaaaga cttgatc 27

<210> 9  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic

<400> 9  
gatcaagtct ttttgcggcc gcg 23

<210> 10  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic

<400> 10  
aatcgcggc cgcaaaaaga cttcg 25

<210> 11  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic

<400> 11  
cgaagtcttt ttgcggccgc g 21

<210> 12  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic

<221> modified\_base  
<222> (0)...(0)  
<223> n=5-bromo-deoxyuridine

<400> 12  
aaaaatnnnn cnaagtcttt nt 22

<210> 13  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic

<221> modified\_base  
<222> (0)...(0)  
<223> n=5-iodo-deoxyuridine

<400> 13  
aaaaatnnnn caaagtcttt tt 22

<210> 14  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic oligonucleotide

<400> 14  
aattcgcgcc cgcaaaaaga cttgatc 27

<210> 15  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic oligonucleotide

<400> 15  
gatcaagtct ttttcgccc gcg 23

<210> 16  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> synthetic oligonucleotide

<400> 16  
aattcgcgcc cgcaaaaaga cttgatcaag tctttttgcg gcgcgcg 46

<210> 17  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 17

aattcgcggc cgcaaaaaga cttgatcaag tctttttgcg gccgcg

46

<210> 18

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 18

aaaaagactt agaaaaatTT tt

22

<210> 19

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic oligonucleotide

<400> 19

aaaaattttt ctaagtcttt tt

22